



Content validity analysis of literacy assessment instruments

Dameis Surya Anggara*, Candra Abdillah

Universitas Pamulang, Indonesia

*Corresponding Author: dameis_surya@yahoo.com

ABSTRACT

This research is part of the development of teacher assessment literacy instruments. The research objective at this stage was to determine the content validity of the instrument. The instruments developed were 25 multiple-choice items constructed from Assessment Objectives, Measurement Theory, Assessment Process, and Assessment Fairness. Then the instrument is assessed qualitatively through a Focus Group Discussion to collect various suggestions and input from various points of view. After that the instrument was assessed quantitatively by 3 experts with 6 aspects of assessment, namely the relevance of the item to the assessment literacy dimension, relevance of items with indicators, clarity of the main items, logical answer options, standardization of language, and the functioning of descriptions/ cases/ pictures/ tables on items. The collected data were analyzed according to the type of data. Qualitative data analyzed using a qualitative approach, while quantitative data were analyzed using Content Validity Index (CVI). The results showed that of the 25 items, each had a CVI value of more than 0.60 with an average of 0.96. So it can be concluded that the items and instruments of assessment literacy have valid criteria based on content.

Keywords: literacy assessment; content validity; instrument

Article history

Received:

04 September 2022

Revised:

12 October 2022

Accepted:

24 January 2023

Published:

19 May 2023

Citation (APA Style): Anggara, D. S. & Abdillah, C. (2023). Content validity analysis of literacy assessment instruments. *Cakrawala Pendidikan: Jurnal Ilmiah Pendidikan*, 42(2), 447-459. DOI: <https://doi.org/10.21831/cp.v42i2.55900>

INTRODUCTION

Assessment is the process of collecting and processing information to measure the achievement of student learning outcomes. The ability to use knowledge when carrying out an assessment is called assessment literacy. Therefore, assessment literacy involves cognitive and skills abilities when carrying out assessments. Cognitive abilities play a role in absorbing knowledge about the concept of assessment, which is obtained when reading. Then the knowledge that has been obtained will be internalized into skilled actions when carrying out the assessment. This is in accordance with the explanation of Brookhart (2011) and Popham (2013) which states that assessment literacy is the skills and knowledge of teachers to measure and support student learning through the implementation of assessments.

Assessment literacy involves the ability to construct reliable assessments and then administer and score these assessments to facilitate valid instructional decisions (Popham, 2013; Stiggins, 2004). Assessment literacy describes the teacher's ability to plan for, administer, understand and apply the outcomes of assessments accurately and efficiently (Boyle 2005; Stiggins 1999; Stoyhoff and Chapelle, 2005). In addition, Assessment literacy can empower the teachers to use data collected from various assessment methods, interpret it properly and make their instruction better (Gotch, 2012).

A teacher must have good assessment literacy. This is because through good assessment, it will help teachers overcome learning problems, measure the effectiveness of the learning process, and control the achievement of learning objectives. This is in accordance with the explanation of Brookhart (2011) and Popham (2013) which states that assessment literacy can measure and

support student learning. Furthermore, Shams & Iqbal (2019) states that assessment literacy has benefits for developing teacher professional competence, improving the welfare of teachers, students, and institutions, helping teachers overcome learning problems in class, and can improve the quality of learning. In addition, assessment literacy aims to develop teacher professional competence, determine the success of student achievement, develop teacher abilities in applying assessment methods according to student character, and develop students' academic abilities and potential (Johnston & Costello, 2005; Zolfaghari & Ahmadi, 2016). All these benefits greatly affect the quality of education.

But behind the importance of the assessment literacy described above, there are various problems faced by teachers. DeLuca, et al (2016) stated that the low understanding and application of teacher assessments was caused by teachers lacking competency in interpreting, integrating, and communicating assessment results. The same thing happened to teachers in Indonesia. Many teachers do not understand the concept of assessment and its function, so that the application is not optimal. When the teacher conducts an assessment, it is only limited to making measurements without being accompanied by a comprehensive interpretation and without communicating the results of the assessment to related parties.

Based on interviews and observations in several schools, many teachers do not know the various ways of learning assessment. They only know that assessing means giving test questions to students, then analyzing them based on the correct answers. In addition, there are teachers who do not understand the steps in developing learning assessments such as making assessment designs, developing items, making scoring and making assessment guidelines. The majority of them directly took questions from other sources without any analysis of the assessment stages. The teacher is also still confused in analyzing the results of the assessment in accordance with the applicable assessment norms. Then if there are description questions, they give scores based on their own assumptions without a standard rubric. Of course, these problems have an impact on the low quality of education. Van de Grift (2007); Levi & Inbar-Lourie (2020) stated that the low literacy assessment of teachers will have an impact on decreasing the quality of teaching, student activity, and learning achievement.

The problems described above must be corrected immediately so that the quality of education is guaranteed and can even increase. One solution is through an assessment literacy diagnostic test which can be used as a basis for making policies to improve weaknesses related to assessment literacy. So that the test results can describe the real situation, a valid and reliable instrument is needed to measure assessment literacy. These instruments can be made through a development process based on standards and dimensions of teacher assessment literacy.

Developing an assessment instrument must be based on the dimensions of the assessment literacy (O'Loughlin, 2013; Taylor, 2009; Koh, 2011). Assessment literacy standards include choosing an assessment method, developing assessment methods, administering, assigning, and interpreting learning outcomes, using assessment outcomes in decision making, using assessment to determine levels of learning outcomes, communicating assessment outcomes, and knowing unethical practices (Yamtim & Wongwanich, 2014). In addition, DeLuca et al (2015) explained that assessment literacy indicators are: 1) assessment objectives which are the process of selecting the appropriate form of assessment based on the instructional objectives of the assessment, for example: diagnostic tests, formative tests, summative tests, and selection tests; 2) measurement theory which is an understanding of the psychometric nature of assessment, for example: validity test and reliability test; 3) the assessment process, which is the ability to compile, manage, evaluate, and interpret the results of the assessment. At this stage, it requires the ability to communicate the objectives, processes and results of the assessment to stakeholders such as students, parents and other parties; and 4) assessment fairness, which is the ability to give fair assessments to students based on the ethics of assessment. The ethics of assessment in question is to disclose accurate information on the results of assessments by protecting students' rights and privacy.

Research on assessment literacy has been carried out by DeLuca, et al (2015) which resulted in an analysis of assessment literacy standards for schoolteachers and university lecturers. And continued again by DeLuca, et al (2016) who produced instruments to support teacher

assessment literacy with a class assessment inventory approach. Furthermore, research conducted by Zolfaghari and Ahmadi (2017) resulted in an analysis of the teacher's assessment of literacy skills obtained through interviews. As well as Ashraf and Zolfaghari (2018) which produced a relationship between EFL Teachers' Assessment Literacy and Their Reflective Teaching. Based on previous studies, there has been no research that has produced an assessment literacy test instrument. The majority of related studies only analyze assessment literacy using a questionnaire instrument, of course, will produce data that is less than ideal, therefore the purpose of this study is to produce an assessment literacy test instrument that has content valid criteria.

METHOD

The purpose of this study was to determine the content validity of the teacher assessment literacy instrument obtained through the product development process. Therefore, the research methodology used is the Research and Development Methodology with limitations on certain steps. Borg and Gall (1984) state that Research and Development (R and D) methodology is a process in developing and validating educational products. In addition, Richey and Client (2007) state that R and D is a systematic study in developing and evaluating the products created.

Thiagarajan (1974) states that R and D consists of four stages, namely define, design, develop, and disseminate. Cohen and Swerdlik (2010) state that the test development process consists of 5 steps, namely conceptualizing the test, constructing the test, testing the test, conducting item analysis, and revising the test. Chadha (2009) states that there are five steps in constructing a test, namely planning the test, preparing the initial draft of the test, trying out the initial draft of the test, evaluating the test, and constructing the final draft of the test. In addition, Irwing and Hughes (2018) state that there are 10 steps in developing a test, namely definition constructs, test specifications, and test structure, overall planning, item development such as construct definition, item creation, item review, testing items, scale construction, reliability, validation, test scoring, specification test, implementation and test, and technical manual.

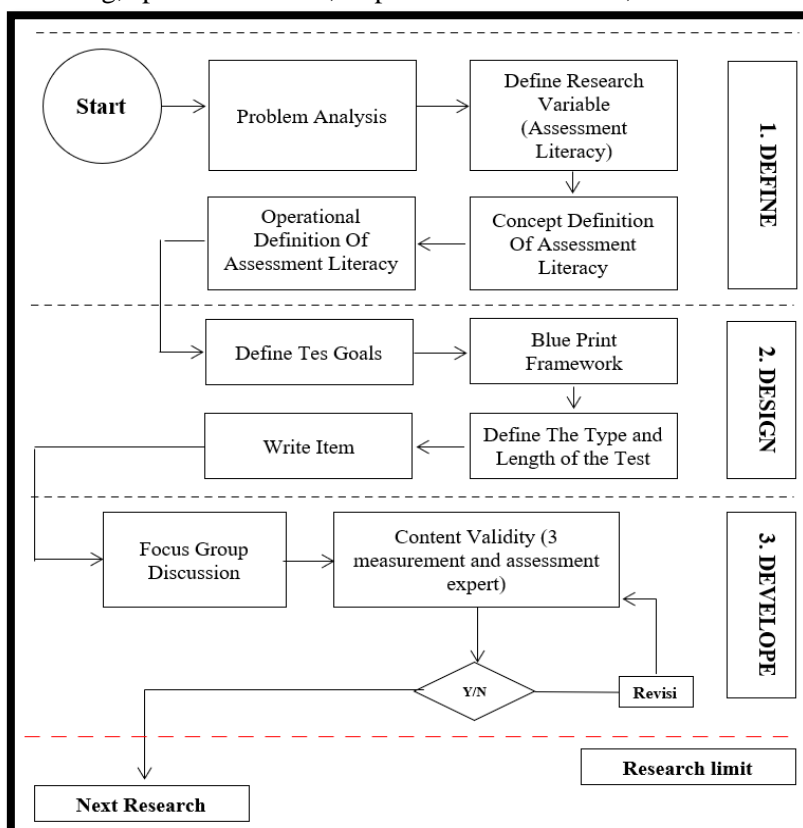


Figure 1. Research Procedure

Based on the description above, this research procedure was designed based on R and D theory and instrument development theory. However, in this study, the R and D steps were limited to the develop stage, namely the analysis of the content validity of the assessment literacy instrument. The procedure steps are presented in Figure 1.

Based on Figure 1 it can be explained that the procedure of this research is *the first*, analyzing the problem. The first step in this research is to analyze the problem. Based on interviews and observations of several teachers, it can be concluded that teacher assessment literacy has low criteria with some of the evidence that has been explained in the introduction section.

The second is determining research variables. After the problem is known, the second step is to determine the research variable, namely the assessment literacy variable. These variables are determined based on the problems found in the field.

The third is defining the literacy assessment concept. After determining the variable, the next step is to define the concept of the variable. The concept of assessment literacy is the skills and knowledge of teachers to measure and support student learning through the implementation of assessments.

The fourth is defining operational literacy assessment. The operational definition of assessment literacy is the teacher's skills and knowledge to measure and support student learning through the implementation of assessments, which are measured through tests with dimensions including assessment objectives, measurement theory, assessment process, and assessment fairness.

The fifth is determining the purpose of the test. The purpose of the test is to diagnose teacher assessment literacy skills, which is used as a basis for making policies in an effort to improve teacher assessment literacy problems.

The sixth is designing blueprint design. Mardapi (2008) stated that the steps for making a blueprint design include writing general objectives, making a list of topics, determining indicators, and determining the number of questions.

The seventh is determining the type and length of the test. The type of test used in this study is a multiple-choice test, with the reason that it can cover a wide range of material and a short processing time. While the length of the test is obtained through an analysis of the estimated time for each question. Nitko (1996) estimates the time to work on analysis-type questions is 2-5 minutes for each question.

The eighth is writing items. Mardapi (2008) states that the main guidelines for making multiple-choice tests include: the subject matter must be clear, the sentences used are appropriate to the developmental level of the test takers, the language used is standard, the location of the correct answer choices is determined randomly, all answer choices are logical, length the answer choice sentences are relatively the same, and there are no clues to the correct answer.

The ninth is focus group discussion. Focus Group Discussion is a qualitative assessment activity of each item of the assessment literacy instrument. Qualitative assessment in the form of repair notes on item construction. Of course, in accordance with the 6 aspects of the assessment that have been described previously. Referring to the notes, the instrument was repaired and given to experts to be assessed quantitatively.

The tenth is analyze content validity. Content validity was determined through an assessment by experts of the instrument. Goodwin and Leech (2003) stated that content-based validity is based on logical analysis and expert evaluation of content measurements such as: subject matter, item format, and constituent sentences. This assessment is a quantitative assessment by giving a score ranging from 1 to 4 for each aspect of the assessment.

The resulting data is then analyzed according to the data type. Qualitative data analyzed using a qualitative approach, while quantitative data were analyzed using Content Validity Index (CVI). The description of the data analysis is presented in table 1.

Table 1. Data Analysis

Types of Data	Data analysis	Description	Criteria
Qualitative	Qualitative Analysis	Used as a guideline for improving draft items	-
Quantitative	Content Validity Index (CVI)	Used as a guideline in determining content validity	$0,60 \leq CVI \leq 1,00$

FINDING AND DISCUSSION

Findings

The literacy assessment instrument developed in this study was used to diagnose the teacher's initial ability to assess literacy in learning. This instrument was developed with a total of 25 items. These items were developed referring to the dimensions of assessment literacy, such as assessment purpose, assessment process, assessment fairness, and measurement theory. The summary is presented in table 2.

Table 2. Teacher Assessment Literacy Instrument Blueprint

No	Dimensions	Indicator	Item Number	Number of Items
1	Assessment Purpose	Analyze the use of this type of assessment based on instructional objectives; Analyze the function of formative assessment in teaching and learning.	1, 2, 3, 4 5, 6	6
2	Assessment Process	Construct, administer, and assess, and interpret assessment results for the basis of instructional decision making; Communicating the purpose, process, and results of the assessment to students, parents/guardians, and other stakeholders	7, 8, 9, 10, 11 12, 13	7
3	Assessment Fairness	Creating fair assessment conditions for all students; Disclose accurate information about the assessment.	14, 15, 16, 17 18, 19	6
4	Measurement Theory	Analyze the nature of validity in the assessment; Analyze the nature of validity in the assessment	20, 21 22, 23, 24, 25	6
			Total	25

Referring to Table 2, then the type and length of the test are determined. The type of test used in this study is a multiple-choice test, with the reason that it can measure a wide range of material and a short processing time. Then the length of the test is determined through an analysis of the estimated processing time for each question presented in table 3.

Table 3. Length of Assessment Literacy Test

No	Dimensions	Number of Items	Estimated Time (Nitko, 1996)	Time Required
1	Assessment Purpose	6	2 – 5 minutes/ item	12 – 30 minutes
2	Assessment Process	7	2 – 5 minutes/ item	14 – 35 minutes
3	Assessment Fairness	6	2 – 5 minutes/ item	12 – 30 minutes
4	Measurement Theory	6	2 – 5 minutes/ item	12 – 30 minutes
Total		25		50 – 125 menit

Based on table 3, it can be explained that 25 multiple choice items can be completed within 50 – 125 minutes. On that basis, the length of the assessment literacy test is 60 minutes.

After the blue print, the type of test, and the length of the test are determined, the next step is to write instrument items. Writing literacy assessment items refers to aspects, such as: 1) the relevance of the item to the quantitative literacy dimension, 2) the relevance of the item to the question indicators, 3) the clarity of the subject matter, 4) the logic of all answer options, 5) the standardization of the language used, and 6) the functioning of case descriptions/discourses/pictures/tables/graphics in the item. Writing this item will produce a draft literacy assessment instrument.

The draft instruments that have been developed are then assessed qualitatively through Focus Group Discussion (FGD) activities, involving teachers and experts. The purpose of a qualitative assessment is to collect various suggestions and input from various points of view. The following is a description of the results of the assessment:

Suggestions and improvements to item number 1

The results of the FGD concluded that item number 1 did not measure the indicator item. The resulting recommendations are items replaced with other items. Evidence of this assessment is presented in Figure 2.

KARTU SOAL PILIHAN GANDA									
<table border="1"> <thead> <tr> <th colspan="2">IDENTITAS</th> </tr> </thead> <tbody> <tr> <td>Bentuk Penilaian</td> <td>: Pilihan Ganda</td> </tr> <tr> <td>Estimasi Waktu</td> <td>: 2 – 5 menit</td> </tr> <tr> <td>Penyusun</td> <td>: Tim Peneliti</td> </tr> </tbody> </table>		IDENTITAS		Bentuk Penilaian	: Pilihan Ganda	Estimasi Waktu	: 2 – 5 menit	Penyusun	: Tim Peneliti
IDENTITAS									
Bentuk Penilaian	: Pilihan Ganda								
Estimasi Waktu	: 2 – 5 menit								
Penyusun	: Tim Peneliti								
<table border="1"> <thead> <tr> <th>KOMPETENSI</th> <th>NO SOAL</th> </tr> </thead> <tbody> <tr> <td>Literasi Penilaian</td> <td>1</td> </tr> </tbody> </table>		KOMPETENSI	NO SOAL	Literasi Penilaian	1				
KOMPETENSI	NO SOAL								
Literasi Penilaian	1								
<table border="1"> <thead> <tr> <th>DIMENSI</th> <th>KUNCI JAWABAN</th> </tr> </thead> <tbody> <tr> <td>Tujuan Penilaian</td> <td>C</td> </tr> </tbody> </table>		DIMENSI	KUNCI JAWABAN	Tujuan Penilaian	C				
DIMENSI	KUNCI JAWABAN								
Tujuan Penilaian	C								
<table border="1"> <thead> <tr> <th colspan="2">INDIKATOR</th> </tr> </thead> <tbody> <tr> <td>Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional</td> <td></td> </tr> </tbody> </table>		INDIKATOR		Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional					
INDIKATOR									
Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional									
<table border="1"> <thead> <tr> <th>SARAN PERBAIKAN</th> </tr> </thead> <tbody> <tr> <td> </td> </tr> </tbody> </table>		SARAN PERBAIKAN							
SARAN PERBAIKAN									
<table border="1"> <thead> <tr> <th>BUTIR SOAL</th> </tr> </thead> <tbody> <tr> <td>Terdapat contoh soal berikut: <div style="border: 1px solid black; padding: 2px;"> Apa yang Anda ketahui tentang "Peristiwa Pertempuran 5 Hari di Semarang" pada Tahun 1945? </div> Contoh soal di atas berbentuk A. Uraian Tertutup B. Uraian Terbatas C. Uraian Terbuka D. Uraian Singkat E. Uraian Terpimpin <i>Uraian soal belum mengukur indikator, diteliti di sesuaikan kembali</i> </td> </tr> </tbody> </table>		BUTIR SOAL	Terdapat contoh soal berikut: <div style="border: 1px solid black; padding: 2px;"> Apa yang Anda ketahui tentang "Peristiwa Pertempuran 5 Hari di Semarang" pada Tahun 1945? </div> Contoh soal di atas berbentuk A. Uraian Tertutup B. Uraian Terbatas C. Uraian Terbuka D. Uraian Singkat E. Uraian Terpimpin <i>Uraian soal belum mengukur indikator, diteliti di sesuaikan kembali</i>						
BUTIR SOAL									
Terdapat contoh soal berikut: <div style="border: 1px solid black; padding: 2px;"> Apa yang Anda ketahui tentang "Peristiwa Pertempuran 5 Hari di Semarang" pada Tahun 1945? </div> Contoh soal di atas berbentuk A. Uraian Tertutup B. Uraian Terbatas C. Uraian Terbuka D. Uraian Singkat E. Uraian Terpimpin <i>Uraian soal belum mengukur indikator, diteliti di sesuaikan kembali</i>									

Figure 2. Suggestions for Improving Item Number 1

Based on Figure 2, it was then revised with the results presented in Figure 3.

KARTU SOAL PILIHAN GANDA									
<table border="1"> <thead> <tr> <th colspan="2">IDENTITAS</th> </tr> </thead> <tbody> <tr> <td>Bentuk Penilaian</td> <td>: Pilihan Ganda</td> </tr> <tr> <td>Estimasi Waktu</td> <td>: 2 – 5 menit</td> </tr> <tr> <td>Penyusun</td> <td>: Tim Peneliti</td> </tr> </tbody> </table>		IDENTITAS		Bentuk Penilaian	: Pilihan Ganda	Estimasi Waktu	: 2 – 5 menit	Penyusun	: Tim Peneliti
IDENTITAS									
Bentuk Penilaian	: Pilihan Ganda								
Estimasi Waktu	: 2 – 5 menit								
Penyusun	: Tim Peneliti								
<table border="1"> <thead> <tr> <th>KOMPETENSI</th> <th>NO SOAL</th> </tr> </thead> <tbody> <tr> <td>Literasi Penilaian</td> <td>3</td> </tr> </tbody> </table>		KOMPETENSI	NO SOAL	Literasi Penilaian	3				
KOMPETENSI	NO SOAL								
Literasi Penilaian	3								
<table border="1"> <thead> <tr> <th>DIMENSI</th> <th>KUNCI JAWABAN</th> </tr> </thead> <tbody> <tr> <td>Tujuan Penilaian</td> <td>E</td> </tr> </tbody> </table>		DIMENSI	KUNCI JAWABAN	Tujuan Penilaian	E				
DIMENSI	KUNCI JAWABAN								
Tujuan Penilaian	E								
<table border="1"> <thead> <tr> <th colspan="2">INDIKATOR</th> </tr> </thead> <tbody> <tr> <td>Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional</td> <td></td> </tr> </tbody> </table>		INDIKATOR		Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional					
INDIKATOR									
Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional									
<table border="1"> <thead> <tr> <th>SARAN PERBAIKAN</th> </tr> </thead> <tbody> <tr> <td> </td> </tr> </tbody> </table>		SARAN PERBAIKAN							
SARAN PERBAIKAN									
<table border="1"> <thead> <tr> <th>BUTIR SOAL</th> </tr> </thead> <tbody> <tr> <td>Dalam suatu diskusi Antoni berpendapat bahwa untuk mengetahui kemampuan original suatu siswa, maka bentuk penilaian yang digunakan adalah Tes Uraian. Namun Arya tidak setuju dengan pendapat tersebut. Jika pendapat Arya dinilai benar, manakah bentuk penilaian yang dimaksud ? A. Pilihan Ganda B. Benar Salah C. Isian Singkat D. Uraian Terbuka E. Praktik </td> </tr> </tbody> </table>		BUTIR SOAL	Dalam suatu diskusi Antoni berpendapat bahwa untuk mengetahui kemampuan original suatu siswa, maka bentuk penilaian yang digunakan adalah Tes Uraian. Namun Arya tidak setuju dengan pendapat tersebut. Jika pendapat Arya dinilai benar, manakah bentuk penilaian yang dimaksud ? A. Pilihan Ganda B. Benar Salah C. Isian Singkat D. Uraian Terbuka E. Praktik						
BUTIR SOAL									
Dalam suatu diskusi Antoni berpendapat bahwa untuk mengetahui kemampuan original suatu siswa, maka bentuk penilaian yang digunakan adalah Tes Uraian. Namun Arya tidak setuju dengan pendapat tersebut. Jika pendapat Arya dinilai benar, manakah bentuk penilaian yang dimaksud ? A. Pilihan Ganda B. Benar Salah C. Isian Singkat D. Uraian Terbuka E. Praktik									

Figure 3. Revision Result of Item Number 1

Suggestions and Improvements to Item Number 2

The results of the FGD concluded that there was no relationship between the questions and the answer options. Questions contain types of assessment, while answer options contain various forms of response. This situation will cause item bias. The resulting recommendations are revised

questions and adapted to the answer options. Evidence of these recommendations is presented in Figure 4.

KARTU SOAL PILIHAN GANDA									
<table border="1"> <thead> <tr> <th colspan="2">IDENTITAS</th> </tr> </thead> <tbody> <tr> <td>Bentuk Penilaian</td> <td>: Pilihan Ganda</td> </tr> <tr> <td>Estimasi Waktu</td> <td>: 2 – 5 menit</td> </tr> <tr> <td>Penyusun</td> <td>: Tim Peneliti</td> </tr> </tbody> </table>		IDENTITAS		Bentuk Penilaian	: Pilihan Ganda	Estimasi Waktu	: 2 – 5 menit	Penyusun	: Tim Peneliti
IDENTITAS									
Bentuk Penilaian	: Pilihan Ganda								
Estimasi Waktu	: 2 – 5 menit								
Penyusun	: Tim Peneliti								
<table border="1"> <thead> <tr> <th>KOMPETENSI</th> <th>NO SOAL</th> </tr> </thead> <tbody> <tr> <td>Literasi Penilaian</td> <td>2</td> </tr> </tbody> </table>		KOMPETENSI	NO SOAL	Literasi Penilaian	2				
KOMPETENSI	NO SOAL								
Literasi Penilaian	2								
<table border="1"> <thead> <tr> <th>DIMENSI</th> <th>KUNCI JAWABAN</th> </tr> </thead> <tbody> <tr> <td>Tujuan Penilaian</td> <td>A</td> </tr> </tbody> </table>		DIMENSI	KUNCI JAWABAN	Tujuan Penilaian	A				
DIMENSI	KUNCI JAWABAN								
Tujuan Penilaian	A								
<table border="1"> <thead> <tr> <th colspan="2">INDIKATOR</th> </tr> </thead> <tbody> <tr> <td>Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional</td> <td></td> </tr> </tbody> </table>		INDIKATOR		Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional					
INDIKATOR									
Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional									
<table border="1"> <thead> <tr> <th colspan="2">SARAN PERBAIKAN</th> </tr> </thead> <tbody> <tr> <td colspan="2"> </td> </tr> </tbody> </table>		SARAN PERBAIKAN							
SARAN PERBAIKAN									
<table border="1"> <thead> <tr> <th colspan="2">BUTIR SOAL</th> </tr> </thead> <tbody> <tr> <td colspan="2">Seorang guru ingin menguji ketrampilan siswa dalam membuat gambar bangun datar. Bentuk penilaian yang cocok untuk mengukur ketrampilan tersebut adalah...</td> </tr> <tr> <td colspan="2"> A. Authentic Response B. Constructive Response C. Selected Response D. True False Response E. Pairing Response </td> </tr> <tr> <td colspan="2"> <i>Kalimat tanya lebih baik menjadi : " Bentuk response yang diharapkan dalam pengerjaan tes adalah "</i> </td> </tr> </tbody> </table>		BUTIR SOAL		Seorang guru ingin menguji ketrampilan siswa dalam membuat gambar bangun datar. Bentuk penilaian yang cocok untuk mengukur ketrampilan tersebut adalah...		A. Authentic Response B. Constructive Response C. Selected Response D. True False Response E. Pairing Response		<i>Kalimat tanya lebih baik menjadi : " Bentuk response yang diharapkan dalam pengerjaan tes adalah "</i>	
BUTIR SOAL									
Seorang guru ingin menguji ketrampilan siswa dalam membuat gambar bangun datar. Bentuk penilaian yang cocok untuk mengukur ketrampilan tersebut adalah...									
A. Authentic Response B. Constructive Response C. Selected Response D. True False Response E. Pairing Response									
<i>Kalimat tanya lebih baik menjadi : " Bentuk response yang diharapkan dalam pengerjaan tes adalah "</i>									

Figure 4. Suggestions for Improving Item Number 2

Based on Figure 4, it was then revised with the results presented in Figure 5.

KARTU SOAL PILIHAN GANDA									
<table border="1"> <thead> <tr> <th colspan="2">IDENTITAS</th> </tr> </thead> <tbody> <tr> <td>Bentuk Penilaian</td> <td>: Pilihan Ganda</td> </tr> <tr> <td>Estimasi Waktu</td> <td>: 2 – 5 menit</td> </tr> <tr> <td>Penyusun</td> <td>: Tim Peneliti</td> </tr> </tbody> </table>		IDENTITAS		Bentuk Penilaian	: Pilihan Ganda	Estimasi Waktu	: 2 – 5 menit	Penyusun	: Tim Peneliti
IDENTITAS									
Bentuk Penilaian	: Pilihan Ganda								
Estimasi Waktu	: 2 – 5 menit								
Penyusun	: Tim Peneliti								
<table border="1"> <thead> <tr> <th>KOMPETENSI</th> <th>NO SOAL</th> </tr> </thead> <tbody> <tr> <td>Literasi Penilaian</td> <td>2</td> </tr> </tbody> </table>		KOMPETENSI	NO SOAL	Literasi Penilaian	2				
KOMPETENSI	NO SOAL								
Literasi Penilaian	2								
<table border="1"> <thead> <tr> <th>DIMENSI</th> <th>KUNCI JAWABAN</th> </tr> </thead> <tbody> <tr> <td>Tujuan Penilaian</td> <td>A</td> </tr> </tbody> </table>		DIMENSI	KUNCI JAWABAN	Tujuan Penilaian	A				
DIMENSI	KUNCI JAWABAN								
Tujuan Penilaian	A								
<table border="1"> <thead> <tr> <th colspan="2">INDIKATOR</th> </tr> </thead> <tbody> <tr> <td>Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional</td> <td></td> </tr> </tbody> </table>		INDIKATOR		Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional					
INDIKATOR									
Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional									
<table border="1"> <thead> <tr> <th colspan="2">SARAN PERBAIKAN</th> </tr> </thead> <tbody> <tr> <td colspan="2"> </td> </tr> </tbody> </table>		SARAN PERBAIKAN							
SARAN PERBAIKAN									
<table border="1"> <thead> <tr> <th colspan="2">BUTIR SOAL</th> </tr> </thead> <tbody> <tr> <td colspan="2">Seorang guru ingin menguji ketrampilan siswa dalam membuat gambar bangun datar. Bentuk respons yang dihasilkan dalam pengujian tersebut adalah...</td> </tr> <tr> <td colspan="2"> A. Authentic Response B. Constructive Response C. Selected Response D. True False Response E. Pairing Response </td> </tr> </tbody> </table>		BUTIR SOAL		Seorang guru ingin menguji ketrampilan siswa dalam membuat gambar bangun datar. Bentuk respons yang dihasilkan dalam pengujian tersebut adalah...		A. Authentic Response B. Constructive Response C. Selected Response D. True False Response E. Pairing Response			
BUTIR SOAL									
Seorang guru ingin menguji ketrampilan siswa dalam membuat gambar bangun datar. Bentuk respons yang dihasilkan dalam pengujian tersebut adalah...									
A. Authentic Response B. Constructive Response C. Selected Response D. True False Response E. Pairing Response									

Figure 5. Revision Result of Item Number 2

Suggestions and Improvements to Item Number 3

The results of the FGD concluded that the main item information was incomplete, causing item bias. The resulting recommendations are to complete the main item information so that there is conformity with the answer options. Evidence of these recommendations is presented in Figure 6.

KARTU SOAL PILIHAN GANDA									
<table border="1"> <thead> <tr> <th colspan="2">IDENTITAS</th> </tr> </thead> <tbody> <tr> <td>Bentuk Penilaian</td> <td>: Pilihan Ganda</td> </tr> <tr> <td>Estimasi Waktu</td> <td>: 2 – 5 menit</td> </tr> <tr> <td>Penyusun</td> <td>: Tim Peneliti</td> </tr> </tbody> </table>		IDENTITAS		Bentuk Penilaian	: Pilihan Ganda	Estimasi Waktu	: 2 – 5 menit	Penyusun	: Tim Peneliti
IDENTITAS									
Bentuk Penilaian	: Pilihan Ganda								
Estimasi Waktu	: 2 – 5 menit								
Penyusun	: Tim Peneliti								
<table border="1"> <thead> <tr> <th>KOMPETENSI</th> <th>NO SOAL</th> </tr> </thead> <tbody> <tr> <td>Literasi Penilaian</td> <td>3</td> </tr> </tbody> </table>		KOMPETENSI	NO SOAL	Literasi Penilaian	3				
KOMPETENSI	NO SOAL								
Literasi Penilaian	3								
<table border="1"> <thead> <tr> <th>DIMENSI</th> <th>KUNCI JAWABAN</th> </tr> </thead> <tbody> <tr> <td>Tujuan Penilaian</td> <td>E</td> </tr> </tbody> </table>		DIMENSI	KUNCI JAWABAN	Tujuan Penilaian	E				
DIMENSI	KUNCI JAWABAN								
Tujuan Penilaian	E								
<table border="1"> <thead> <tr> <th colspan="2">INDIKATOR</th> </tr> </thead> <tbody> <tr> <td>Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional</td> <td></td> </tr> </tbody> </table>		INDIKATOR		Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional					
INDIKATOR									
Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional									
<table border="1"> <thead> <tr> <th colspan="2">SARAN PERBAIKAN</th> </tr> </thead> <tbody> <tr> <td colspan="2"> </td> </tr> </tbody> </table>		SARAN PERBAIKAN							
SARAN PERBAIKAN									
<table border="1"> <thead> <tr> <th colspan="2">BUTIR SOAL</th> </tr> </thead> <tbody> <tr> <td colspan="2">Dalam suatu diskusi Antoni berpendapat bahwa untuk mengetahui kemampuan original suatu siswa, maka bentuk penilaian yang digunakan adalah Tes Uraian. Namun Arya tidak setuju dengan pendapat tersebut. Manakah bentuk penilaian yang dimaksud oleh Arya ?</td> </tr> <tr> <td colspan="2"> A. Pilihan Ganda B. Benar Salah C. Isian Singkat D. Uraian Terbuka E. Praktik </td> </tr> <tr> <td colspan="2"> <i>agar pertanyaaan tidak bias, maka dirumuskan menjadi berikut : " Jika pendapat Arya bernilai benar, manakah bentuk penilaian yang dimaksud ? "</i> </td> </tr> </tbody> </table>		BUTIR SOAL		Dalam suatu diskusi Antoni berpendapat bahwa untuk mengetahui kemampuan original suatu siswa, maka bentuk penilaian yang digunakan adalah Tes Uraian. Namun Arya tidak setuju dengan pendapat tersebut. Manakah bentuk penilaian yang dimaksud oleh Arya ?		A. Pilihan Ganda B. Benar Salah C. Isian Singkat D. Uraian Terbuka E. Praktik		<i>agar pertanyaaan tidak bias, maka dirumuskan menjadi berikut : " Jika pendapat Arya bernilai benar, manakah bentuk penilaian yang dimaksud ? "</i>	
BUTIR SOAL									
Dalam suatu diskusi Antoni berpendapat bahwa untuk mengetahui kemampuan original suatu siswa, maka bentuk penilaian yang digunakan adalah Tes Uraian. Namun Arya tidak setuju dengan pendapat tersebut. Manakah bentuk penilaian yang dimaksud oleh Arya ?									
A. Pilihan Ganda B. Benar Salah C. Isian Singkat D. Uraian Terbuka E. Praktik									
<i>agar pertanyaaan tidak bias, maka dirumuskan menjadi berikut : " Jika pendapat Arya bernilai benar, manakah bentuk penilaian yang dimaksud ? "</i>									

Figure 6. Suggestions for Improving Item Number 3

Based on Figure 6, it was then revised with the results presented in Figure 7.

KARTU SOAL PILIHAN GANDA	
IDENTITAS	
Bentuk Penilaian	: Pilihan Ganda
Estimasi Waktu	: 2 – 5 menit
Penyusun	: Tim Peneliti
KOMPETENSI	NO SOAL
Literasi Penilaian	3
DIMENSI	KUNCI JAWABAN
Tujuan Penilaian	E
INDIKATOR	
Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional	
SARAN PERBAIKAN	
BUTIR SOAL	
Dalam suatu diskusi Antoni berpendapat bahwa untuk mengetahui kemampuan original suatu siswa, maka bentuk penilaian yang digunakan adalah Tes Uraian. Namun Arya tidak setuju dengan pendapat tersebut. Jika pendapat Arya dinilai benar, manakah bentuk penilaian yang dimaksud ?	
A. Pilihan Ganda B. Benar Salah C. Isian Singkat D. Uraian Terbuka E. Praktik	

Figure 7. Revision Result of Item Number 3

Suggestions and Improvements to Item Number 5

The results of the FGD concluded that the symbol for inequality in answer options B and C needs to be replaced. Evidence of recommendations is presented in Figure 8.

KARTU SOAL PILIHAN GANDA																																																																																																																																																																																					
IDENTITAS																																																																																																																																																																																					
Bentuk Penilaian	: Pilihan Ganda																																																																																																																																																																																				
Estimasi Waktu	: 2 – 5 menit																																																																																																																																																																																				
Penyusun	: Tim Peneliti																																																																																																																																																																																				
KOMPETENSI	NO SOAL																																																																																																																																																																																				
Literasi Penilaian	5																																																																																																																																																																																				
DIMENSI	KUNCI JAWABAN																																																																																																																																																																																				
Tujuan Penilaian	A																																																																																																																																																																																				
INDIKATOR																																																																																																																																																																																					
Menganalisis fungsi penilaian formatif dalam pengajaran dan pembelajaran																																																																																																																																																																																					
SARAN PERBAIKAN																																																																																																																																																																																					
BUTIR SOAL																																																																																																																																																																																					
Perhatikan hasil tes Kompetensi Dasar 3.5 pada mata pelajaran IPS terhadap 12 siswa di SMP Cendekia:																																																																																																																																																																																					
<table border="1"> <thead> <tr> <th rowspan="2">No</th> <th rowspan="2">Kode Siswa</th> <th colspan="10">Nomor Soal/ Skor</th> <th rowspan="2">Total</th> </tr> <tr> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> <th>7</th> <th>8</th> <th>9</th> <th>10</th> </tr> </thead> <tbody> <tr><td>1</td><td>A</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>3</td></tr> <tr><td>2</td><td>B</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>9</td></tr> <tr><td>3</td><td>C</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>7</td></tr> <tr><td>4</td><td>D</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>8</td></tr> <tr><td>5</td><td>E</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>6</td></tr> <tr><td>6</td><td>F</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>9</td></tr> <tr><td>7</td><td>G</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>8</td></tr> <tr><td>8</td><td>H</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>6</td></tr> <tr><td>9</td><td>I</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>2</td></tr> <tr><td>10</td><td>J</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>3</td></tr> <tr><td>11</td><td>K</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>7</td></tr> <tr><td>12</td><td>L</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>7</td></tr> <tr><td colspan="2">Jumlah</td><td>10</td><td>2</td><td>4</td><td>8</td><td>9</td><td>8</td><td>9</td><td>11</td><td>11</td><td>3</td></tr> </tbody> </table>	No	Kode Siswa	Nomor Soal/ Skor										Total	1	2	3	4	5	6	7	8	9	10	1	A	1	0	0	1	1	0	0	0	0	0	3	2	B	1	1	1	1	1	1	1	1	0	9	3	C	1	0	0	1	1	1	1	1	0	7	4	D	1	0	0	1	1	1	1	1	1	8	5	E	1	1	0	0	0	1	0	1	1	6	6	F	1	0	1	1	1	1	1	1	1	9	7	G	1	0	1	1	1	1	1	1	0	8	8	H	1	0	0	0	1	1	1	1	0	6	9	I	0	0	0	0	0	0	1	1	0	2	10	J	0	0	0	0	0	0	1	1	0	3	11	K	1	0	1	1	1	0	1	1	0	7	12	L	1	0	0	1	1	1	1	1	0	7	Jumlah		10	2	4	8	9	8	9	11	11	3	<p>Jika hasil tes formatif yang Anda lakukan hasilnya seperti pada Tabel di atas, apa langkah yang harus Anda lakukan ?</p> <p>A. Terdapat 3 butir soal yaitu butir 2, 3, dan 10 yang mayoritas belum dapat dicapai oleh siswa, sehingga perlu dilakukan perbaikan pembelajaran;</p> <p>B. Terdapat 3 siswa yang hanya mencapai $\leq 25\%$ dari tujuan pembelajaran, sehingga perlu dianalisis kesulitan belajar siswa;</p> <p>C. Terdapat 9 siswa yang dapat mencapai $\leq 60\%$ dari tujuan pembelajaran, sehingga dapat dilanjutkan ke Kompetensi Dasar berikutnya;</p> <p>D. Butir soal tersulit adalah nomor 2, sehingga perlu dipertimbangkan kembali untuk diujikan;</p> <p>E. Butir soal termudah adalah nomor 8 dan 9, sehingga perlu dipertimbangkan kembali untuk diujikan;</p>
No			Kode Siswa	Nomor Soal/ Skor										Total																																																																																																																																																																							
	1	2		3	4	5	6	7	8	9	10																																																																																																																																																																										
1	A	1	0	0	1	1	0	0	0	0	0	3																																																																																																																																																																									
2	B	1	1	1	1	1	1	1	1	0	9																																																																																																																																																																										
3	C	1	0	0	1	1	1	1	1	0	7																																																																																																																																																																										
4	D	1	0	0	1	1	1	1	1	1	8																																																																																																																																																																										
5	E	1	1	0	0	0	1	0	1	1	6																																																																																																																																																																										
6	F	1	0	1	1	1	1	1	1	1	9																																																																																																																																																																										
7	G	1	0	1	1	1	1	1	1	0	8																																																																																																																																																																										
8	H	1	0	0	0	1	1	1	1	0	6																																																																																																																																																																										
9	I	0	0	0	0	0	0	1	1	0	2																																																																																																																																																																										
10	J	0	0	0	0	0	0	1	1	0	3																																																																																																																																																																										
11	K	1	0	1	1	1	0	1	1	0	7																																																																																																																																																																										
12	L	1	0	0	1	1	1	1	1	0	7																																																																																																																																																																										
Jumlah		10	2	4	8	9	8	9	11	11	3																																																																																																																																																																										

Figure 8. Suggestions for Improving Item Number 5

KARTU SOAL PILIHAN GANDA																																																																																																																																																																																				
IDENTITAS																																																																																																																																																																																				
Bentuk Penilaian	: Pilihan Ganda																																																																																																																																																																																			
Estimasi Waktu	: 2 – 5 menit																																																																																																																																																																																			
Penyusun	: Tim Peneliti																																																																																																																																																																																			
KOMPETENSI	NO SOAL																																																																																																																																																																																			
Literasi Penilaian	5																																																																																																																																																																																			
DIMENSI	KUNCI JAWABAN																																																																																																																																																																																			
Tujuan Penilaian	A																																																																																																																																																																																			
INDIKATOR																																																																																																																																																																																				
Menganalisis fungsi penilaian formatif dalam pengajaran dan pembelajaran																																																																																																																																																																																				
SARAN PERBAIKAN																																																																																																																																																																																				
BUTIR SOAL																																																																																																																																																																																				
Perhatikan hasil tes Kompetensi Dasar 3.5 pada mata pelajaran IPS terhadap 12 siswa di SMP Cendekia:																																																																																																																																																																																				
<table border="1"> <thead> <tr> <th rowspan="2">No</th> <th rowspan="2">Kode Siswa</th> <th colspan="10">Nomor Soal/ Skor</th> <th rowspan="2">Total</th> </tr> <tr> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> <th>7</th> <th>8</th> <th>9</th> <th>10</th> </tr> </thead> <tbody> <tr><td>1</td><td>A</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>3</td></tr> <tr><td>2</td><td>B</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>9</td></tr> <tr><td>3</td><td>C</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>7</td></tr> <tr><td>4</td><td>D</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>8</td></tr> <tr><td>5</td><td>E</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>6</td></tr> <tr><td>6</td><td>F</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>9</td></tr> <tr><td>7</td><td>G</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>8</td></tr> <tr><td>8</td><td>H</td><td>1</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>6</td></tr> <tr><td>9</td><td>I</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>2</td></tr> <tr><td>10</td><td>J</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>0</td><td>3</td></tr> <tr><td>11</td><td>K</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>7</td></tr> <tr><td>12</td><td>L</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>1</td><td>0</td><td>7</td></tr> <tr><td colspan="2">Jumlah</td><td>10</td><td>2</td><td>4</td><td>8</td><td>9</td><td>8</td><td>9</td><td>11</td><td>11</td><td>3</td></tr> </tbody> </table>	No	Kode Siswa	Nomor Soal/ Skor										Total	1	2	3	4	5	6	7	8	9	10	1	A	1	0	0	1	1	0	0	0	0	3	2	B	1	1	1	1	1	1	1	1	0	9	3	C	1	0	0	1	1	1	1	1	0	7	4	D	1	0	0	1	1	1	1	1	1	8	5	E	1	1	0	0	0	1	0	1	1	6	6	F	1	0	1	1	1	1	1	1	1	9	7	G	1	0	1	1	1	1	1	1	0	8	8	H	1	0	0	0	1	1	1	1	0	6	9	I	0	0	0	0	0	0	1	1	0	2	10	J	0	0	0	0	0	0	1	1	0	3	11	K	1	0	1	1	1	0	1	1	0	7	12	L	1	0	0	1	1	1	1	1	0	7	Jumlah		10	2	4	8	9	8	9	11	11	3	<p>Jika hasil tes formatif yang Anda lakukan hasilnya seperti pada tabel di atas, apa langkah yang harus Anda lakukan ?</p> <p>A. Terdapat 3 butir soal yaitu butir 2, 3, dan 10 yang mayoritas belum dapat dicapai oleh siswa, sehingga perlu dilakukan perbaikan pembelajaran;</p> <p>B. Terdapat 3 siswa yang hanya mencapai $\leq 25\%$ dari tujuan pembelajaran, sehingga perlu dianalisis kesulitan belajar siswa;</p> <p>C. Terdapat 9 siswa yang dapat mencapai $\leq 60\%$ dari tujuan pembelajaran, sehingga dapat dilanjutkan ke Kompetensi Dasar berikutnya;</p> <p>D. Butir soal tersulit adalah nomor 2, sehingga perlu dipertimbangkan kembali untuk diujikan;</p> <p>E. Butir soal termudah adalah nomor 8 dan 9, sehingga perlu dipertimbangkan kembali untuk diujikan;</p>
No			Kode Siswa	Nomor Soal/ Skor										Total																																																																																																																																																																						
	1	2		3	4	5	6	7	8	9	10																																																																																																																																																																									
1	A	1	0	0	1	1	0	0	0	0	3																																																																																																																																																																									
2	B	1	1	1	1	1	1	1	1	0	9																																																																																																																																																																									
3	C	1	0	0	1	1	1	1	1	0	7																																																																																																																																																																									
4	D	1	0	0	1	1	1	1	1	1	8																																																																																																																																																																									
5	E	1	1	0	0	0	1	0	1	1	6																																																																																																																																																																									
6	F	1	0	1	1	1	1	1	1	1	9																																																																																																																																																																									
7	G	1	0	1	1	1	1	1	1	0	8																																																																																																																																																																									
8	H	1	0	0	0	1	1	1	1	0	6																																																																																																																																																																									
9	I	0	0	0	0	0	0	1	1	0	2																																																																																																																																																																									
10	J	0	0	0	0	0	0	1	1	0	3																																																																																																																																																																									
11	K	1	0	1	1	1	0	1	1	0	7																																																																																																																																																																									
12	L	1	0	0	1	1	1	1	1	0	7																																																																																																																																																																									
Jumlah		10	2	4	8	9	8	9	11	11	3																																																																																																																																																																									

Figure 9. Revision Result of Item Number 5

Suggestions and Improvements to Item Number 9

The results of the FGD concluded that answer options D and E in item number 9 were correct. The resulting recommendation is that one of the answer options is changed to be false according to Figure 10.

KARTU SOAL PILIHAN GANDA									
<table border="1"> <tr><th colspan="2">IDENTITAS</th></tr> <tr><td>Bentuk Penilaian</td><td>: Pilihan Ganda</td></tr> <tr><td>Estimasi Waktu</td><td>: 2 – 5 menit</td></tr> <tr><td>Penyusun</td><td>: Tim Peneliti</td></tr> </table>	IDENTITAS		Bentuk Penilaian	: Pilihan Ganda	Estimasi Waktu	: 2 – 5 menit	Penyusun	: Tim Peneliti	<p>Andi sedang melakukan penelitian tugas akhir di kelas X. Setelah memberikan soal tes pilihan ganda ke 30 siswa, kemudian dia memberikan skor dengan kriteria jawaban benar mendapatkan skor 2 dan jawaban salah mendapatkan skor 1. Apakah yang dilakukan Andi sudah tepat ?</p> <p>A. Sudah tepat, karena skor benar lebih banyak daripada skor salah B. Sudah tepat, karena skor interval antara 1 sampai dengan 2 C. Sudah tepat, karena masing-masing jawaban memiliki skor yang telah ditentukan D. Belum tepat, karena skor salah seharusnya $\frac{1}{2}$ E. Belum tepat, karena skor benar harusnya 1 dan skor salah adalah 0</p> <p><i>Pilihan jawaban "0" juga dinilai benar. Lebih baik diganti agar jawaban bernilai salah.</i></p>
IDENTITAS									
Bentuk Penilaian	: Pilihan Ganda								
Estimasi Waktu	: 2 – 5 menit								
Penyusun	: Tim Peneliti								
<table border="1"> <tr><th>KOMPETENSI</th><th>NO SOAL</th></tr> <tr><td>Literasi Penilaian</td><td>9</td></tr> </table>	KOMPETENSI	NO SOAL	Literasi Penilaian	9					
KOMPETENSI	NO SOAL								
Literasi Penilaian	9								
<table border="1"> <tr><th>DIMENSI</th><th>KUNCI JAWABAN</th></tr> <tr><td>Tujuan Penilaian</td><td>E</td></tr> </table>	DIMENSI	KUNCI JAWABAN	Tujuan Penilaian	E					
DIMENSI	KUNCI JAWABAN								
Tujuan Penilaian	E								
<table border="1"> <tr><th colspan="2">INDIKATOR</th></tr> <tr><td colspan="2">Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional</td></tr> </table>	INDIKATOR		Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional						
INDIKATOR									
Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional									
<table border="1"> <tr><th>SARAN PERBAIKAN</th></tr> <tr><td> </td></tr> </table>	SARAN PERBAIKAN								
SARAN PERBAIKAN									

Figure 10. Suggestions for Improving Item Number 9

Based on Figure 10, it was then revised with the results presented in Figure 11.

KARTU SOAL PILIHAN GANDA									
<table border="1"> <tr><th colspan="2">IDENTITAS</th></tr> <tr><td>Bentuk Penilaian</td><td>: Pilihan Ganda</td></tr> <tr><td>Estimasi Waktu</td><td>: 2 – 5 menit</td></tr> <tr><td>Penyusun</td><td>: Tim Peneliti</td></tr> </table>	IDENTITAS		Bentuk Penilaian	: Pilihan Ganda	Estimasi Waktu	: 2 – 5 menit	Penyusun	: Tim Peneliti	<p>Andi sedang melakukan penelitian tugas akhir di kelas X. Setelah memberikan soal tes pilihan ganda ke 30 siswa, kemudian dia memberikan skor dengan kriteria jawaban benar mendapatkan skor 2 dan jawaban salah mendapatkan skor 1. Apakah yang dilakukan Andi sudah tepat ?</p> <p>A. Sudah tepat, karena skor benar lebih banyak daripada skor salah B. Sudah tepat, karena skor interval antara 1 sampai dengan 2 C. Sudah tepat, karena masing-masing jawaban memiliki skor yang telah ditentukan D. Belum tepat, karena skor salah seharusnya $\frac{1}{2}$ E. Belum tepat, karena skor benar harusnya 1 dan skor salah adalah 0</p>
IDENTITAS									
Bentuk Penilaian	: Pilihan Ganda								
Estimasi Waktu	: 2 – 5 menit								
Penyusun	: Tim Peneliti								
<table border="1"> <tr><th>KOMPETENSI</th><th>NO SOAL</th></tr> <tr><td>Literasi Penilaian</td><td>9</td></tr> </table>	KOMPETENSI	NO SOAL	Literasi Penilaian	9					
KOMPETENSI	NO SOAL								
Literasi Penilaian	9								
<table border="1"> <tr><th>DIMENSI</th><th>KUNCI JAWABAN</th></tr> <tr><td>Proses Penilaian</td><td>E</td></tr> </table>	DIMENSI	KUNCI JAWABAN	Proses Penilaian	E					
DIMENSI	KUNCI JAWABAN								
Proses Penilaian	E								
<table border="1"> <tr><th colspan="2">INDIKATOR</th></tr> <tr><td colspan="2">Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional</td></tr> </table>	INDIKATOR		Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional						
INDIKATOR									
Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional									
<table border="1"> <tr><th>SARAN PERBAIKAN</th></tr> <tr><td> </td></tr> </table>	SARAN PERBAIKAN								
SARAN PERBAIKAN									

Figure 11. Revision Result of Item Number 9

Suggestions and Improvements to Item Number 10

The results of the FGD concluded that the information on the main items was incomplete, so that there were no correct answer options. The resulting recommendations are the subject matter added to the information in accordance with Figure 12.

KARTU SOAL PILIHAN GANDA									
<table border="1"> <tr><th colspan="2">IDENTITAS</th></tr> <tr><td>Bentuk Penilaian</td><td>: Pilihan Ganda</td></tr> <tr><td>Estimasi Waktu</td><td>: 2 – 5 menit</td></tr> <tr><td>Penyusun</td><td>: Tim Peneliti</td></tr> </table>	IDENTITAS		Bentuk Penilaian	: Pilihan Ganda	Estimasi Waktu	: 2 – 5 menit	Penyusun	: Tim Peneliti	<p>Siswa kelas V yang bernama Ari menjawab 10 soal pilihan ganda dengan tepat dari 15 soal yang diujikan. Skor yang diperoleh adalah 5. Ilustrasi tersebut menunjukkan pemberian skor berjenis....</p> <p>A. Skoring Tanpa Denda B. Skoring Dengan Denda C. Skoring Rubrik D. Skoring Dengan Pembobotan E. Skoring Tanpa Pembobotan</p> <p><i>Pokok soal ditambahkan informasi "jawaban yg salah", agar pertanyaannya tidak bias. Jika dengan denda 10 benar - 5 salah = skor 5</i></p>
IDENTITAS									
Bentuk Penilaian	: Pilihan Ganda								
Estimasi Waktu	: 2 – 5 menit								
Penyusun	: Tim Peneliti								
<table border="1"> <tr><th>KOMPETENSI</th><th>NO SOAL</th></tr> <tr><td>Literasi Penilaian</td><td>10</td></tr> </table>	KOMPETENSI	NO SOAL	Literasi Penilaian	10					
KOMPETENSI	NO SOAL								
Literasi Penilaian	10								
<table border="1"> <tr><th>DIMENSI</th><th>KUNCI JAWABAN</th></tr> <tr><td>Tujuan Penilaian</td><td>E B</td></tr> </table>	DIMENSI	KUNCI JAWABAN	Tujuan Penilaian	E B					
DIMENSI	KUNCI JAWABAN								
Tujuan Penilaian	E B								
<table border="1"> <tr><th colspan="2">INDIKATOR</th></tr> <tr><td colspan="2">Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional</td></tr> </table>	INDIKATOR		Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional						
INDIKATOR									
Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional									
<table border="1"> <tr><th>SARAN PERBAIKAN</th></tr> <tr><td>Jawaban menjadi B. Jenis scoring untuk bentuk penilaian pilhan ganda, ex: 10 benar - 5 salah = 5.</td></tr> </table>	SARAN PERBAIKAN	Jawaban menjadi B. Jenis scoring untuk bentuk penilaian pilhan ganda, ex: 10 benar - 5 salah = 5.							
SARAN PERBAIKAN									
Jawaban menjadi B. Jenis scoring untuk bentuk penilaian pilhan ganda, ex: 10 benar - 5 salah = 5.									

Figure 12. Suggestions for Improving Item Number 10

Based on Figure 12, it was then revised with the results presented in Figure 13.

KARTU SOAL PILIHAN GANDA	
IDENTITAS	
Bentuk Penilaian : Pilihan Ganda	
Estimasi Waktu : 2 – 5 menit	
Penyusun : Tim Peneliti	
KOMPETENSI	NO SOAL
Literasi Penilaian	10
DIMENSI	KUNCI JAWABAN
Proses Penilaian	B
INDIKATOR	
Membangun, mengelola, dan menilai, dan menafsirkan hasil penilaian untuk dasar pengambilan keputusan instruksional	
SARAN PERBAIKAN	
BUTIR SOAL	
Siswa kelas V yang bernama Ari menjawab 10 soal pilihan ganda dengan tepat dan menjawab 5 soal dengan salah dari 15 soal yang diujikan. Skor yang diperoleh adalah 5. Ilustrasi tersebut menunjukkan pemberian skor berjenis....	
A. Skoring Tanpa Denda	
B. Skoring Dengan Denda	
C. Skoring Rubrik	
D. Skoring Dengan Pembobotan	
E. Skoring Tanpa Pembobotan	

Figure 13. Revision Result of Item Number 10

After the instrument has been assessed qualitatively and revised, the next step is to provide a draft of the instrument to 3 experts to be assessed quantitatively. The purpose of this quantitative assessment is to determine the content validity of the teacher's assessment literacy instrument. Quantitative assessment is carried out by giving a score between 1 to 4 on each aspect of the assessment based on the question card. These assessment aspects include the relevance of the item to the assessment literacy dimension, the relevance of items with indicators, clarity of the main items, logical answer options, standardization of language, and the functioning of descriptions/ cases/ pictures/ tables on items. While the question card guidelines are presented in Figure 14.

KARTU SOAL PILIHAN GANDA			
IDENTITAS			
Bentuk Penilaian : Pilihan Ganda			
Estimasi Waktu : 2 – 5 menit			
Penyusun : Tim Peneliti			
KOMPETENSI	NO SOAL		
Literasi Penilaian	3		
DIMENSI	KUNCI JAWABAN		
Tujuan Penilaian	E		
INDIKATOR			
Menganalisis penggunaan bentuk penilaian berdasarkan tujuan instruksional			
SARAN PERBAIKAN			
BUTIR SOAL			
Dalam suatu diskusi Antoni berpendapat bahwa untuk mengetahui kemampuan original suatu siswa, maka bentuk penilaian yang digunakan adalah Tes Uraian. Namun Arya tidak setuju dengan pendapat tersebut. Jika pendapat Arya dinilai benar, manakah bentuk penilaian yang dimaksud ?			
A. Pilihan Ganda			
B. Benar Salah			
C. Isian Singkat			
D. Uraian Terbuka			
E. Praktik			
PENILAIAN			
No	Aspek Penilaian	Skala Penilaian	
1	Relevansi butir soal dengan dimensi literasi kuantitatif	Tidak Relevan	① ② ③ ④ Sangat Relevan
2	Relevansi butir soal dengan indikator soal	Tidak Relevan	① ② ③ ④ Sangat Relevan
3	Kejelasan pokok soal	Tidak Jelas	① ② ③ ④ Sangat Jelas
4	Kelogisan pada semua opsi jawaban	Tidak Logis	① ② ③ ④ Sangat Logis
5	Keterbakuan bahasa yang digunakan	Tidak Baku	① ② ③ ④ Sangat Baku
6	Keberfungsian uraian kasus/wacana/ gambar/ tabel/ grafik dalam butir soal	Tidak Berfungsi	① ② ③ ④ Sangat Berfungsi

Figure 14. Example of Question Cards

The results of the quantitative assessment are presented in table 4.

Table 4. Content validity index results

Item Number	Proportion of CVI/Expert Items			Mean I-CVI	Conclusion
	Expert 1	Expert 1	Expert 1		
1	1	1	1	1.00	Valid Content
2	1	1	1	1.00	Valid Content
3	1	1	0.83	0.94	Valid Content
4	1	1	1	1.00	Valid Content
5	1	1	0.67	0.89	Valid Content
6	1	1	1	1.00	Valid Content
7	1	1	1	1.00	Valid Content

8	1	1	1	1.00	Valid Content
9	1	1	1	1.00	Valid Content
10	1	1	1	1.00	Valid Content
11	1	0.67	1	0.89	Valid Content
12	1	1	0.83	0.94	Valid Content
13	1	1	1	1.00	Valid Content
14	1	1	1	1.00	Valid Content
15	1	1	1	1.00	Valid Content
16	0.67	1	1	0.89	Valid Content
17	1	0.67	1	0.89	Valid Content
18	1	0.83	1	0.94	Valid Content
19	1	1	1	1.00	Valid Content
20	0.67	1	1	0.89	Valid Content
21	1	1	1	1.00	Valid Content
22	1	1	1	1.00	Valid Content
23	1	0.83	1	0.94	Valid Content
24	1	1	0.67	0.89	Valid Content
25	1	1	1	1.00	Valid Content
Mean CVI				0.96	Valid Content

Based on table 4, it can be explained that all items have a CVI value of more than 0.60, it can be concluded that all literacy assessment instrument items have content valid criteria. Then, if seen from the average CVI value, it has a value of $0.96 > 0.60$, it can be concluded that the assessment literacy instrument also has content valid criteria. This means that the items that have been developed in this study have a construct that is in accordance with the dimensions and indicators of assessment literacy.

Discussion

Based on the research findings, the literacy assessment instrument developed was of the multiple-choice type with a total of 25 items and a test length of 60 minutes. These items were developed based on the construct of assessment literacy, which consists of 4 dimensions, namely the purpose of the assessment, the theory of measurement, the process of assessment, and the fairness of the assessment. In addition, items are also developed based on measuring indicators for each dimension, clarity of main items, logical answers, standardization of language, and the functioning of item descriptions such as cases/discourses/images/tables/ or graphs. DeLuca et al (2015) state that the dimensions of assessment literacy are assessment purpose, assessment process, assessment fairness, and measurement theory.

The selection of multiple-choice tests is based on the breadth of assessment literacy variables consisting of 4 dimensions and 8 item indicators. In order to be done in a relatively short time, a multiple-choice test was selected. This is in accordance with the opinion of Cohen and Swerdlik (2010) which states that this type of multiple-choice test can measure a broad range of material with a relatively short processing time. Then the length of the test is 60 minutes based on the estimated length of the test by Nitko (1996) which is presented in table 3.

Then the target number of items must have at least 12 items or 3 items for each dimension with valid and reliable criteria. Therefore, the questions are developed at least twice or more than the targeted questions, namely as many as 25 items. Neill (2011) which states that one domain has at least 3 items. Sumintono and Widhiarso (2015) stated that the number of items developed by researchers must be two or three times the target number of items, because if there are items that do not pass the selection, there are still other reserve items.

Items are developed according to the assessment literacy dimensions and item structure, such as: item clarity, logical answers, language standardization, and the functioning of question descriptions such as cases/discourses, pictures/tables/graphs. DeLuca et al (2015) state that the indicators of assessment literacy are the purpose of the assessment; assessment process; fairness of judgment; and measurement theory. Mardapi (2008) states that the main guidelines for making multiple choice tests include: the subject matter must be clear, the sentences used are appropriate

to the level of development of the test takers, the language used is standard, the location of the correct answer choices is determined randomly, all answer choices are logical, the sentence length of the answer choices is relatively the same, there is no hint of the correct answer.

Instrument draft, then analyzed for content validity by 3 experts. Goodwin and Leech (2003) state that content-based validity is based on logical analysis and expert evaluation of item constructs, such as: item subject matter, item format, and constituent sentences. Furthermore, the item construct as described above, refers to the opinion of Mardapi (2008: 93) which states that the main guidelines in making multiple-choice tests include: the subject matter must be clear, the sentences used are appropriate to the level of development of the test takers, the language used is standard, the location of the answer choices correct answer is determined randomly, all answer choices are logical, the sentence length of the answer choices is relatively the same, there are no clues to the correct answer. The number of experts is in accordance with the opinion of Lynn (1986) which states that the number of experts used in expert validation is at least 3 experts and no more than 10 experts.

Then the expert's assessment data were analyzed using the CVI (Content Validity Index) with the result that all items had an I-CVI value ≥ 0.60 , so all items had content valid criteria. As for the overall CVI score, it is $0.96 \geq 0.60$, so the instrument has content-valid criteria. Polit and Beck (2006) state that CVI is a calculation of the proportion of items that get a score of 3 or 4 from experts. Rempusheski and O'Hara (2005) state that the recommended proportion of CVI ranges from 0.60 to 1.0.

CONCLUSION

Based on the research results, it can be explained that the value of the Content Validity Index (CVI) for each item of the quantitative literacy instrument exceeds 0.60. Meanwhile, the Content Validity Index (CVI) value for the literacy assessment instrument is 0.96 which exceeds 0.60. So it can be concluded that the items and literacy assessment instruments that have been developed have valid criteria in terms of content. This means that the constructs of the items and instruments are in accordance with the theoretical constructs used.

REFERENCES

- Anggara, D. S., & Haribowo, H. (2021). Content validity analysis on mathematical quantitative literacy instruments in elementary schools. *Journal of Physics: Conference Series*, 1918 (4). <https://doi.org/10.1088/1742-6596/1918/4/042091>.
- Ashraf, H., & Zolfaghari, S. (2018). EFL teachers' assessment literacy and their reflective teaching. *International Journal of Instruction*, 11(1), 425-436.
- Borg, W.R. dan Gall, M.D. (1984). *Educational Research: An Introduction*. Longman.
- Boyle-Baise, M. (2005). Preparing community-oriented teachers: Reflections from a multicultural service-learning project. *Journal of Teacher Education*, 56(5), 446-458. <https://doi.org/10.1177/0022487105282113>
- Brookhart, S. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30, 3–12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Chadha, N. K. (2009). *Applied psychometry*. SAGE Publications India. <https://doi.org/10.4135/9788132108221>
- Cohen, R., & Swerdlik, M. (2010). *Psychological testing and assessment: An introduction to tests and measurement 7th*. McGraw-Hill Primis.
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Approaches to classroom assessment inventory: A new instrument to support teacher assessment literacy. *Educational Assessment*, 21(4), 248-266. <http://dx.doi.org/10.1080/10627197.2016.1236677>
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28, 251–272.

- Goodwin, L.D., & Leech, N.L. (2003). The meaning of validity in the new standards for educational and psychological testing. *Measurement and Evaluation in Counseling and Development*, 36(3), 181-191. <https://doi.org/10.1080/07481756.2003.11909741>
- Gotch, C. M. (2012). An investigation of teacher educational measurement literacy. A *Dissertation*. Washington State University, Department of Educational Leadership and Counseling Psychology
- Irwing, P., Booth, T., & Hughes, D. J. (Eds.). (2018). *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development*. John Wiley & Sons. <https://doi.org/10.1002/9781118489772>
- Johnston, P., & Costello, P. (2005). Principles for literacy assessment. *Reading research quarterly*, 40(2), 256-267. <https://doi.org/10.1598/RRQ.40.2.6>
- Koh, K., Burke, L. E. C.-A., Luke, A., Gong, W., & Tan, C. (2017). Developing the assessment literacy of teachers in Chinese language classrooms: A focus on assessment task design. *Language Teaching Research*, 22(3), 264-288. <https://doi.org/10.1177/1362168816684366>
- Levi, T., & Inbar-Lourie, O. (2020). Assessment literacy or language assessment literacy: Learning from the teachers. *Language Assessment Quarterly*, 17(2), 168-182. <https://doi.org/10.1080/15434303.2019.1692347>
- Lynn, M.R. (1986). Determination and Quantification of Content Validity. *Nursing Research*, 35, 382– 385.
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan non tes*. Mitra Cendekia.
- Polit, D.F. & Beck, C.T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 2006, 29, 489–497. <https://doi.org/10.1002/nur.20147>
- O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing*, 30(3), 363-380. <https://doi.org/10.1177/0265532213480336>
- Popham, W. J. (2013). *Classroom assessment: What teachers need to know 7th*. Pearson
- Rempusheski, V.F. & O'Hara, C.T. (2005). Psychometric properties of the grandparent perceptions of family scale. *Nursing Research*, 54, 419– 427
- Richey, R.C. & Klein. (2007). *Design and Development Research*. Lawrence Erlbaum Associates. Inc.
- Shams, J. A., & Iqbal, M. Z. (2019). Development of classroom assessment literacy training program for university teachers in Punjab. *Bulletin of Education and Research*, 41(2), 41-52.
- Stoynoff, S., & Chapelle, C. A. (2005). *ESOL tests and testing*. TESOL.
- Stiggins, R. J. (1999). Assessment, student confidence, and school success. *The Phi Delta Kappan*, 81(3), 191-198.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86, 22–27.
- Neill, J. (2011). Exploratory factor analysis survey research and design psikologi. *Lecture Material*. Centre or Applied Psychology University of Canberra.
- Nitko, A.J. (1996). *Educational assessment of student 2nd*. Merrill.
- Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29, 21-36. <https://doi.org/10.1017/S0267190509090035>
- Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional Development for Training Teachers of Exceptional Children*. Indiana University Bloomington
- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational research*, 49(2), 127-152. <https://doi.org/10.1080/00131880701369651>
- Yamtim, V., & Wongwanich, S. (2014). A study of classroom assessment literacy of primary school teachers. *Procedia-Social and Behavioral Sciences*, 116, 2998-3004. <https://doi:10.1016/j.sbspro.2014.01.696>
- Zolfaghari, F., & Ahmadi, A. (2016). Assessment literacy components across subject matters. *Cogent Education*, 3(1), 1252561. <https://doi.org/10.1080/2331186X.2016.1252561>